

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/249,003

DATE: 04/28/1999
 TIME: 14:55:13

INPUT SET: S31648.raw

This Raw Listing contains the General
 Information Section and up to the first 5 pages.

SEQUENCE LISTING

ENTERED

(1) General Information:

(i) APPLICANT: Wilson, Peter J
 Morris, Charles P
 Anson, Donald S
 Occhiodoro, Teresa
 Bielicki, Julie
 Clements, Peter R
 Hopwood, John J

(ii) TITLE OF INVENTION: GLYCOSYLATION VARIANTS OF
 IDURONATE 2-SULFATASE

(iii) NUMBER OF SEQUENCES: 15

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Scully, Scott, Murphy & Presser
 (B) STREET: 400 Garden City Plaza
 (C) CITY: Garden City
 (D) STATE: New York
 (E) COUNTRY: USA
 (F) ZIP: 11530

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
 (B) COMPUTER: IBM PC compatible
 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 (D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: 09/249,003
 (B) FILING DATE:
 (C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 07/991,973
 (B) FILING DATE: 17-DEC-1992

(viii) ATTORNEY/AGENT INFORMATION:

RAW SEQUENCE LISTING PATENT APPLICATION US/09/249,003

DATE: 04/28/1999
TIME: 14:55:13

INPUT SET: S31648.raw

```

47
48      (A) NAME: DiGiglio, Frank S
49      (B) REGISTRATION NUMBER: 31,346
50      (C) REFERENCE/DOCKET NUMBER: 8416Z
51
52      (ix) TELECOMMUNICATION INFORMATION:
53
54      (A) TELEPHONE: 516-742-4343
55      (B) TELEFAX: 516-742-4366
56      (C) TELEX: 230 901 SANS UR
57
58      (2) INFORMATION FOR SEQ ID NO:1:
59
60      (i) SEQUENCE CHARACTERISTICS:
61
62      (A) LENGTH: 2297 base pairs
63      (B) TYPE: nucleic acid
64      (C) STRANDEDNESS: double
65      (D) TOPOLOGY: linear
66
67      (ii) MOLECULE TYPE: cDNA
68
69      (ix) FEATURE:
70
71      (A) NAME/KEY: CDS
72      (B) LOCATION: 125..1774
73
74      (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
75
76      CGGCTGTGTT GCGCAGTCTT CATGGGTTC CGACGAGGAG GTCTCTGTGG CTGCGGCGGC      60
77
78      TGCTAACTGC GCCACCTGCT GCAGCCTGTC CCCGCCGCTC TGAAGCGGCC GCGTCGAAGC      120
79
80      CGAA ATG CCG CCA CCC CGG ACC GGC CGA GGC CTT CTC TGG CTG GGT CTG      169
81      Met Pro Pro Pro Arg Thr Gly Arg Gly Leu Leu Trp Leu Gly Leu
82      1          5          10          15
83
84      GTT CTG AGC TCC GTC TGC GTC GCC CTC GGA TCC GAA ACG CAG GCC AAC      217
85      Val Leu Ser Ser Val Cys Val Ala Leu Gly Ser Glu Thr Gln Ala Asn
86      20          25          30
87
88      TCG ACC ACA GAT GCT CTG AAC GTT CTT CTC ATC ATC GTG GAT GAC CTG      265
89      Ser Thr Thr Asp Ala Leu Asn Val Leu Leu Ile Ile Val Asp Asp Leu
90      35          40          45
91
92      CGC CCC TCC CTG GGC TGT TAT GGG GAT AAG CTG GTG AGG TCC CCA AAT      313
93      Arg Pro Ser Leu Gly Cys Tyr Gly Asp Lys Leu Val Arg Ser Pro Asn
94      50          55          60
95
96      ATT GAC CAA CTG GCA TCC CAC AGC CTC CTC TTC CAG AAT GCC TTT GCG      361
97      Ile Asp Gln Leu Ala Ser His Ser Leu Leu Phe Gln Asn Ala Phe Ala
98      65          70          75
99

```

RAW SEQUENCE LISTING PATENT APPLICATION US/09/249,003

DATE: 04/28/1999
TIME: 14:55:14

INPUT SET: S31648.raw

100	CAG CAA GCA GTG TGC GCC CCG AGC CGC GTT TCT TTC CTC ACT GGC AGG	409
101	Gln Gln Ala Val Cys Ala Pro Ser Arg Val Ser Phe Leu Thr Gly Arg	
102	80 85 90 95	
103		
104	AGA CCT GAC ACC ACC CGC CTG TAC GAC TTC AAC TCC TAC TGG AGG GTG	457
105	Arg Pro Asp Thr Thr Arg Leu Tyr Asp Phe Asn Ser Tyr Trp Arg Val	
106	100 105 110	
107		
108	CAC GCT GGA AAC TTC TCC ACC ATC CCC CAG TAC TTC AAG GAG AAT GGC	505
109	His Ala Gly Asn Phe Ser Thr Ile Pro Gln Tyr Phe Lys Glu Asn Gly	
110	115 120 125	
111		
112	TAT GTG ACC ATG TCG GTG GGA AAA GTC TTT CAC CCT GGG ATA TCT TCT	553
113	Tyr Val Thr Met Ser Val Gly Lys Val Phe His Pro Gly Ile Ser Ser	
114	130 135 140	
115		
116	AAC CAT ACC GAT GAT TCT CCG TAT AGC TGG TCT TTT CCA CCT TAT CAT	601
117	Asn His Thr Asp Asp Ser Pro Tyr Ser Trp Ser Phe Pro Pro Tyr His	
118	145 150 155	
119		
120	CCT TCC TCT GAG AAG TAT GAA AAC ACT AAG ACA TGT CGA GGG CCA GAT	649
121	Pro Ser Ser Glu Lys Tyr Glu Asn Thr Lys Thr Cys Arg Gly Pro Asp	
122	160 165 170 175	
123		
124	GGA GAA CTC CAT GCC AAC CTG CTT TGC CCT GTG GAT GTG CTG GAT GTT	697
125	Gly Glu Leu His Ala Asn Leu Leu Cys Pro Val Asp Val Leu Asp Val	
126	180 185 190	
127		
128	CCC GAG GGC ACC TTG CCT GAC AAA CAG AGC ACT GAG CAA GCC ATA CAG	745
129	Pro Glu Gly Thr Leu Pro Asp Lys Gln Ser Thr Glu Gln Ala Ile Gln	
130	195 200 205	
131		
132	TTG TTG GAA AAG ATG AAA ACG TCA GCC AGT CCT TTC TTC CTG GCC GTT	793
133	Leu Leu Glu Lys Met Lys Thr Ser Ala Ser Pro Phe Phe Leu Ala Val	
134	210 215 220	
135		
136	GGG TAT CAT AAG CCA CAC ATC CCC TTC AGA TAC CCC AAG GAA TTT CAG	841
137	Gly Tyr His Lys Pro His Ile Pro Phe Arg Tyr Pro Lys Glu Phe Gln	
138	225 230 235	
139		
140	AAG TTG TAT CCC TTG GAG AAC ATC ACC CTG GCC CCC GAT CCC GAG GTC	889
141	Lys Leu Tyr Pro Leu Glu Asn Ile Thr Leu Ala Pro Asp Pro Glu Val	
142	240 245 250 255	
143		
144	CCT GAT GGC CTA CCC CCT GTG GCC TAC AAC CCC TGG ATG GAC ATC AGG	937
145	Pro Asp Gly Leu Pro Pro Val Ala Tyr Asn Pro Trp Met Asp Ile Arg	
146	260 265 270	
147		
148	CAA CGG GAA GAC GTC CAA GCC TTA AAC ATC AGT GTG CCG TAT GGT CCA	985
149	Gln Arg Glu Asp Val Gln Ala Leu Asn Ile Ser Val Pro Tyr Gly Pro	
150	275 280 285	
151		
152	ATT CCT GTG GAC TTT CAG CGG AAA ATC CGC CAG AGC TAC TTT GCC TCT	1033

RAW SEQUENCE LISTING PATENT APPLICATION US/09/249,003

DATE: 04/28/1999
TIME: 14:55:14

INPUT SET: S31648.raw

153	Ile	Pro	Val	Asp	Phe	Gln	Arg	Lys	Ile	Arg	Gln	Ser	Tyr	Phe	Ala	Ser	
154			290					295					300				
155																	
156	GTG	TCA	TAT	TTG	GAT	ACA	CAG	GTC	GGC	CGC	CTC	TTG	AGT	GCT	TTG	GAC	1081
157	Val	Ser	Tyr	Leu	Asp	Thr	Gln	Val	Gly	Arg	Leu	Leu	Ser	Ala	Leu	Asp	
158		305					310					315					
159																	
160	GAT	CTT	CAG	CTG	GCC	AAC	AGC	ACC	ATC	ATT	GCA	TTT	ACC	TCG	GAT	CAT	1129
161	Asp	Leu	Gln	Leu	Ala	Asn	Ser	Thr	Ile	Ile	Ala	Phe	Thr	Ser	Asp	His	
162	320					325					330					335	
163																	
164	GGG	TGG	GCT	CTA	GGT	GAA	CAT	GGA	GAA	TGG	GCC	AAA	TAC	AGC	AAT	TTT	1177
165	Gly	Trp	Ala	Leu	Gly	Glu	His	Gly	Glu	Trp	Ala	Lys	Tyr	Ser	Asn	Phe	
166					340					345					350		
167																	
168	GAT	GTT	GCT	ACC	CAT	GTT	CCC	CTG	ATA	TTC	TAT	GTT	CCT	GGA	AGG	ACG	1225
169	Asp	Val	Ala	Thr	His	Val	Pro	Leu	Ile	Phe	Tyr	Val	Pro	Gly	Arg	Thr	
170				355					360					365			
171																	
172	GCT	TCA	CTT	CCG	GAG	GCA	GGC	GAG	AAG	CTT	TTC	CCT	TAC	CTC	GAC	CCT	1273
173	Ala	Ser	Leu	Pro	Glu	Ala	Gly	Glu	Lys	Leu	Phe	Pro	Tyr	Leu	Asp	Pro	
174			370					375					380				
175																	
176	TTT	GAT	TCC	GCC	TCA	CAG	TTG	ATG	GAG	CCA	GGC	AGG	CAA	TCC	ATG	GAC	1321
177	Phe	Asp	Ser	Ala	Ser	Gln	Leu	Met	Glu	Pro	Gly	Arg	Gln	Ser	Met	Asp	
178		385					390					395					
179																	
180	CTT	GTG	GAA	CTT	GTG	TCT	CTT	TTT	CCC	ACG	CTG	GCT	GGA	CTT	GCA	GGA	1369
181	Leu	Val	Glu	Leu	Val	Ser	Leu	Phe	Pro	Thr	Leu	Ala	Gly	Leu	Ala	Gly	
182	400					405					410					415	
183																	
184	CTG	CAG	GTT	CCA	CCT	CGC	TGC	CCC	GTT	CCT	TCA	TTT	CAC	GTT	GAG	CTG	1417
185	Leu	Gln	Val	Pro	Pro	Arg	Cys	Pro	Val	Pro	Ser	Phe	His	Val	Glu	Leu	
186					420					425					430		
187																	
188	TGC	AGA	GAA	GGC	AAG	AAC	CTT	CTG	AAG	CAT	TTT	CGA	TTC	CGT	GAC	TTG	1465
189	Cys	Arg	Glu	Gly	Lys	Asn	Leu	Leu	Lys	His	Phe	Arg	Phe	Arg	Asp	Leu	
190				435				440						445			
191																	
192	GAA	GAG	GAT	CCG	TAC	CTC	CCT	GGT	AAT	CCC	CGT	GAA	CTG	ATT	GCC	TAT	1513
193	Glu	Glu	Asp	Pro	Tyr	Leu	Pro	Gly	Asn	Pro	Arg	Glu	Leu	Ile	Ala	Tyr	
194			450					455					460				
195																	
196	AGC	CAG	TAT	CCC	CGG	CCT	TCA	GAC	ATC	CCT	CAG	TGG	AAT	TCT	GAC	AAG	1561
197	Ser	Gln	Tyr	Pro	Arg	Pro	Ser	Asp	Ile	Pro	Gln	Trp	Asn	Ser	Asp	Lys	
198		465					470					475					
199																	
200	CCG	AGT	TTA	AAA	GAT	ATA	AAG	ATC	ATG	GGC	TAT	TCC	ATA	CGC	ACC	ATA	1609
201	Pro	Ser	Leu	Lys	Asp	Ile	Lys	Ile	Met	Gly	Tyr	Ser	Ile	Arg	Thr	Ile	
202	480					485					490					495	
203																	
204	GAC	TAT	AGG	TAT	ACT	GTG	TGG	GTT	GGC	TTC	AAT	CCT	GAT	GAA	TTT	CTA	1657
205	Asp	Tyr	Arg	Tyr	Thr	Val	Trp	Val	Gly	Phe	Asn	Pro	Asp	Glu	Phe	Leu	

RAW SEQUENCE LISTING PATENT APPLICATION US/09/249,003

DATE: 04/28/1999
TIME: 14:55:14

INPUT SET: S31648.raw

	500	505	510	
206				
207				
208	GCT AAC TTT TCT GAC ATC CAT GCA GGG GAA CTG TAT TTT GTG GAT TCT			1705
209	Ala Asn Phe Ser Asp Ile His Ala Gly Glu Leu Tyr Phe Val Asp Ser			
210	515	520	525	
211				
212	GAC CCA TTG CAG GAT CAC AAT ATG TAT AAT GAT TCC CAA GGT GGA GAT			1753
213	Asp Pro Leu Gln Asp His Asn Met Tyr Asn Asp Ser Gln Gly Gly Asp			
214	530	535	540	
215				
216	CTT TTC CAG TTG TTG ATG CCT TGAGTTTTGC CAACCATGGA TGGCAAATGT			1804
217	Leu Phe Gln Leu Leu Met Pro			
218	545	550		
219				
220	GATGTGCTCC CTTCCAGCTG GTGAGAGGAG GAGTTAGAGC TGGTCGTTTT GTGATTACCC			1864
221				
222	ATAATATTGG AAGCAGCCTG AGGGCTAGTT AATCCAAACA TGCATCAACA ATTTGGCCTG			1924
223				
224	AGAATATGTA ACAGCCAAAC CTTTTCGTTT AGTCTTTATT AAAATTTATA ATTGGTAATT			1984
225				
226	GGACCAGTTT TTTTTTTAAT TTCCCTCTTT TAAAACAGT TACGGCTTAT TTAAGAATA			2044
227				
228	AATACAAAGC AAACAAACTC AAGTTATGTC ATACCTTTGG ATACGAAGAC CATAATAAT			2104
229				
230	AACCAAACAT AACATTATAC ACAAAGAATA CTTTCATTAT TTGTGGAATT TAGTGCATTT			2164
231				
232	CAAAAAGTAA TCATATATCA AACTAGGCAC CACACTAAGT TCCTGATTAT TTTGTTTATA			2224
233				
234	ATTTAATAAT ATATCTTATG AGCCCTATAT ATTCAAAATA TTATGTTAAC ATGTAATCCA			2284
235				
236	TGTTTCTTTT TCC			2297
237				
238	(2) INFORMATION FOR SEQ ID NO:2:			
239				
240	(i) SEQUENCE CHARACTERISTICS:			
241				
242	(A) LENGTH: 550 amino acids			
243	(B) TYPE: amino acid			
244	(D) TOPOLOGY: linear			
245				
246	(ii) MOLECULE TYPE: protein			
247				
248	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:			
249				
250	Met Pro Pro Pro Arg Thr Gly Arg Gly Leu Leu Trp Leu Gly Leu Val			
251	1 5 10 15			
252				
253	Leu Ser Ser Val Cys Val Ala Leu Gly Ser Glu Thr Gln Ala Asn Ser			
254	20 25 30			
255				
256	Thr Thr Asp Ala Leu Asn Val Leu Leu Ile Ile Val Asp Asp Leu Arg			
257	35 40 45			
258				

PAGE: 1

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION US/09/249,003

DATE: 04/28/1999
TIME: 14:55:15

INPUT SET: S31648.raw

Line	Error	Original Text
------	-------	---------------